

UC Berkeley

Proposals from the Script Encoding Initiative

Title

Exploratory proposal to encode Germanicist, Nordicist, and other phonetic characters in the UCS

Permalink

<https://escholarship.org/uc/item/4j47b6r7>

Author

Everson, Michael

Publication Date

2008-11-27

Peer reviewed

Universal Multiple-Octet Coded Character Set
International Organization for Standardization
Organisation Internationale de Normalisation
Международная организация по стандартизации

Doc Type: Working Group Document
Title: Exploratory proposal to encode Germanicist, Nordicist,
and other phonetic characters in the UCS
Source: Michael Everson
Status: Individual Contribution
Date: 2008-11-27

This proposal explores the encoding of a fairly large collection of Latin characters for specialist phonetic use, in particular the characters used in the specialized dialect alphabets used in Germany and Austria, in Denmark and Norway, and in Sweden, some additional characters used in older Uralicist publications, as well as some characters used in Sir James Pitmann's Initial Teaching Alphabet, and some phonetic characters devised by Benjamin Franklin.

13 characters tentatively placed in the range 1DE7..1DF3 are combining Latin letters used in German dialectology, in a system called Teuthonista.

2 characters placed in the range 1DFB..1DFC are used with triplets in German dialectology. The triplets can contrast with doublets linked with the DOUBLE BREVE BELOW and DOUBLE INVERTED BREVE. These are distinctive also in legacy font sets used by Germanists.

21 characters placed in a new block at A9E0..A9F4 include 6 parenthesized diacritics above, 1 parenthesized diacritic above, and 7 parenthesized diacritics below. Parenthesization indicates weakened affect of the modifier on pronunciation.

16 characters placed in the SMP at 1E000..1E00F are used in Danish and Norwegian dialectology, in alphabets called "Dania" and "Norvegia" respectively. The reverse italic letters are conventionally rendered vertically when set in italic type. Thus roman type "oq" is intended to be rendered in italic type as "oq". (A similar roman/italic alternation is known in Uralic Phonetic Alphabet practice, where roman "aa" is italicized "aa".)

10 characters placed at 1E010..1E019 for support of early Uralicist phonetic text.

42 characters placed at 1E020..1E049 for support of German dialectology. The first 20 of these are best encoded as units without decomposition; two of them are also found as superscript letters (1DEB, 1DF1).


























































106 characters placed at 1E050..1E0B9 are used in Swedish dialectology, in an alphabet called "Landsmålsalfabetet" in Swedish and sometimes "Suedia" internationally. A good many subtle distinctions including voiced/voiceless pairs are used in this orthography. Some characters are also used in Icelandic dialectology.

18 characters placed at 1E0C0..1E0D1 for support of characters used in Sir James Pitman's Initial Teaching Alphabet.

4 characters placed at 1E0E0..1E0E3 for support of phonetic characters devised by Benjamin Franklin.

9 modifier letter characters placed at 1E0F0..1E0F8 for used in Teuthonista and Norvegia.

Acknowledgements. This project was made possible in part by a grant from the U.S. National Endowment for the Humanities, which funded the which funded the Universal Scripts Project (part of the Script Encoding Initiative at UC Berkeley) in respect of the encoding of phonetic characters for dialectology. Any views, findings, conclusions or recommendations expressed in this publication do not necessarily reflect those of the National Endowment of the Humanities.

	1DC	1DD	1DE	1DF
0	 1DC0	 1DD0	 1DE0	 1DF0
1	 1DC1	 1DD1	 1DE1	 1DF1
2	 1DC2	 1DD2	 1DE2	 1DF2
3	 1DC3	 1DD3	 1DE3	 1DF3
4	 1DC4	 1DD4	 1DE4	
5	 1DC5	 1DD5	 1DE5	
6	 1DC6	 1DD6	 1DE6	
7	 1DC7	 1DD7	 1DE7	
8	 1DC8	 1DD8	 1DE8	
9	 1DC9	 1DD9	 1DE9	
A	 1DCA	 1DDA	 1DEA	
B	 1DCB	 1ddb	 1DEB	 1DFB
C	 1DCC	 1DDC	 1DEC	 1DFC
D	 1DCD	 1DDD	 1DED	 1DFD
E	 1DCE	 1DDE	 1DEE	 1DFE
F	 1DCF	 1DDF	 1DEF	 1DFF

Used for Ancient Greek

These are used as editorial signs for Ancient Greek to indicate scribal deletion of erroneous accent marks.

- 1DC0 ◌̣ COMBINING DOTTED GRAVE ACCENT
→ 1FED ◌̣ greek dialytika and varia
- 1DC1 ◌̣ COMBINING DOTTED ACUTE ACCENT
→ 0344 ◌̣ combining greek dialytika tonos
→ 1FEE ◌̣ greek dialytika and oxia

Miscellaneous marks

- 1DC2 ◌̣ COMBINING SNAKE BELOW
- 1DC3 ◌̣ COMBINING SUSPENSION MARK
Glagolitic
→ 0306 ◌̣ combining breve

Contour tone marks

- 1DC4 ◌̣ COMBINING MACRON-ACUTE
- 1DC5 ◌̣ COMBINING GRAVE-MACRON
- 1DC6 ◌̣ COMBINING MACRON-GRAVE
- 1DC7 ◌̣ COMBINING ACUTE-MACRON
- 1DC8 ◌̣ COMBINING GRAVE-ACUTE-GRAVE
- 1DC9 ◌̣ COMBINING ACUTE-GRAVE-ACUTE

Miscellaneous mark

- 1DCA ◌̣ COMBINING LATIN SMALL LETTER R BELOW

Contour tone marks

- 1DCB ◌̣ COMBINING BREVE-MACRON
Lithuanian dialectology
- 1DCC ◌̣ COMBINING MACRON-BREVE
Lithuanian dialectology

Double diacritic

- 1DCD ◌̣ COMBINING DOUBLE CIRCUMFLEX ABOVE

Medievalist additions

- 1DCE ◌̣ COMBINING OGONEK ABOVE
- 1DCF ◌̣ COMBINING ZIGZAG BELOW
- 1DD0 ◌̣ COMBINING IS BELOW
- 1DD1 ◌̣ COMBINING UR ABOVE
- 1DD2 ◌̣ COMBINING US ABOVE

Medieval superscript letter diacritics

- 1DD3 ◌̣ COMBINING LATIN SMALL LETTER FLATTENED OPEN A ABOVE
- 1DD4 ◌̣ COMBINING LATIN SMALL LETTER AE
- 1DD5 ◌̣ COMBINING LATIN SMALL LETTER AO
- 1DD6 ◌̣ COMBINING LATIN SMALL LETTER AV
- 1DD7 ◌̣ COMBINING LATIN SMALL LETTER C CEDILLA
- 1DD8 ◌̣ COMBINING LATIN SMALL LETTER INSULAR D
- 1DD9 ◌̣ COMBINING LATIN SMALL LETTER ETH
- 1DDA ◌̣ COMBINING LATIN SMALL LETTER G
- 1ddb ◌̣ COMBINING LATIN LETTER SMALL CAPITAL G
- 1DDC ◌̣ COMBINING LATIN SMALL LETTER K
- 1DDD ◌̣ COMBINING LATIN SMALL LETTER L

- 1DDE ◌̣ COMBINING LATIN LETTER SMALL CAPITAL L
- 1DDF ◌̣ COMBINING LATIN LETTER SMALL CAPITAL M
- 1DE0 ◌̣ COMBINING LATIN SMALL LETTER N
- 1DE1 ◌̣ COMBINING LATIN LETTER SMALL CAPITAL N
- 1DE2 ◌̣ COMBINING LATIN LETTER SMALL CAPITAL R
- 1DE3 ◌̣ COMBINING LATIN SMALL LETTER R ROTUNDA
- 1DE4 ◌̣ COMBINING LATIN SMALL LETTER S
- 1DE5 ◌̣ COMBINING LATIN SMALL LETTER LONG S
- 1DE6 ◌̣ COMBINING LATIN SMALL LETTER Z

Superscript letter diacritics for German dialectology

- 1DE7 ◌̣ COMBINING LATIN SMALL LETTER B
- 1DE8 ◌̣ COMBINING LATIN SMALL LETTER SCHWA
- 1DE9 ◌̣ COMBINING LATIN SMALL LETTER F
- 1DEA ◌̣ COMBINING LATIN SMALL LETTER L WITH DOUBLE MIDDLE TILDE
- 1DEB ◌̣ COMBINING LATIN SMALL LETTER LIGHTLY CENTRALIZED O
- 1DEC ◌̣ COMBINING LATIN SMALL LETTER P
- 1DED ◌̣ COMBINING LATIN SMALL LETTER TURNED R
- 1DEE ◌̣ COMBINING LATIN SMALL LETTER SHARP S
- 1DEF ◌̣ COMBINING LATIN SMALL LETTER ESH
- 1DF0 ◌̣ COMBINING LATIN SMALL LETTER UPSILON
- 1DF1 ◌̣ COMBINING LATIN SMALL LETTER LIGHTLY CENTRALIZED U
- 1DF2 ◌̣ COMBINING LATIN SMALL LETTER W
- 1DF3 ◌̣ COMBINING LATIN SMALL LETTER GLOTTAL STOP

Triple diacritics for German dialectology






















- 1DFB ◌̣ COMBINING TRIPLE INVERTED BREVE
→ 0361 ◌̣ combining double inverted breve
- 1DFC ◌̣ COMBINING TRIPLE BREVE BELOW
→ 035C ◌̣ combining double breve below

Additional mark






















- 1DFD ◌̣ COMBINING ALMOST EQUAL TO BELOW

Additional marks for the Uralic Phonetic Alphabet

- 1DFE ◌̣ COMBINING LEFT ARROWHEAD ABOVE
- 1DFF ◌̣ COMBINING RIGHT ARROWHEAD AND DOWN ARROWHEAD BELOW

	A9E	A9F
0	 A9E0	 A9F0
1	 A9E1	 A9F1
2	 A9E2	 A9F2
3	 A9E3	 A9F3
4	 A9E4	 A9F4
5	 A9E5	
6	 A9E6	
7	 A9E7	
8	 A9E8	
9	 A9E9	
A	 A9EA	
B	 A9EB	
C	 A9EC	
D	 A9ED	
E	 A9EE	
F	 A9EF	

Used for German dialectology

A9E0		COMBINING DOUBLED CIRCUMFLEX ACCENT
A9E1		COMBINING DIAERESIS-RING
A9E2		COMBINING PARENTHESIZED TILDE
A9E3		COMBINING DOUBLE PARENTHESIZED TILDE
A9E4		COMBINING PARENTHESIZED MACRON
A9E5		COMBINING PARENTHESIZED BREVE
A9E6		COMBINING PARENTHESIZED DIAERESIS
A9E7		COMBINING PARENTHESIZED CARON
A9E8		COMBINING PARENTHESIZED LATIN SMALL LETTER A
A9E9		COMBINING X-X BELOW
A9EA		COMBINING RING-X BELOW
A9EB		COMBINING DIAERESIS-MACRON-DIAERESIS BELOW
A9EC		COMBINING LENIS MARK BELOW
A9ED		COMBINING DOUBLE LENIS MARK BELOW
A9EE		COMBINING PARENTHESIZED LENIS MARK BELOW
A9EF		COMBINING PARENTHESIZED DOUBLE LENIS MARK BELOW
A9F0		COMBINING PARENTHESIZED DIAERESIS BELOW
A9F1		COMBINING PARENTHESIZED DOT BELOW
A9F2		COMBINING PARENTHESIZED RING BELOW
A9F3		COMBINING PARENTHESIZED VERTICAL LINE BELOW
A9F4		COMBINING PARENTHESIZED INVERTED BREVE BELOW

	1E00	1E01	1E02	1E03	1E04	1E05	1E06	1E07	1E08	1E09	1E0A	1E0B	1E0C	1E0D	1E0E	1E0F
0	à	á	â	ã	ä	å	æ	ç	ð	é	ê	ë	ì	í	î	ï
1	ê	ë	ë	î	ï	ð	é	ê	ë	ì	í	î	ï	ð	é	ê
2	ð	é	ê	ë	ì	í	î	ï	ð	é	ê	ë	ì	í	î	ï
3	é	ê	ë	ì	í	î	ï	ð	é	ê	ë	ì	í	î	ï	ð
4	î	ï	ð	é	ê	ë	ì	í	î	ï	ð	é	ê	ë	ì	í
5	ï	ð	é	ê	ë	ì	í	î	ï	ð	é	ê	ë	ì	í	î
6	ð	é	ê	ë	ì	í	î	ï	ð	é	ê	ë	ì	í	î	ï
7	é	ê	ë	ì	í	î	ï	ð	é	ê	ë	ì	í	î	ï	ð
8	ê	ë	ë	î	ï	ð	é	ê	ë	ì	í	î	ï	ð	é	ê
9	ë	ì	í	î	ï	ð	é	ê	ë	ì	í	î	ï	ð	é	ê
A	ì	í	î	ï	ð	é	ê	ë	ì	í	î	ï	ð	é	ê	ë
B	í	î	ï	ð	é	ê	ë	ì	í	î	ï	ð	é	ê	ë	ì
C	î	ï	ð	é	ê	ë	ì	í	î	ï	ð	é	ê	ë	ì	í
D	ï	ð	é	ê	ë	ì	í	î	ï	ð	é	ê	ë	ì	í	î
E	ð	é	ê	ë	ì	í	î	ï	ð	é	ê	ë	ì	í	î	ï
F	é	ê	ë	ì	í	î	ï	ð	é	ê	ë	ì	í	î	ï	ð

Letters for Danish and Norwegian dialectology

1E000	a	LATIN SMALL LETTER REVERSE ITALIC A
1E001	b	LATIN SMALL LETTER REVERSE ITALIC B
1E002	d	LATIN SMALL LETTER REVERSE ITALIC D
1E003	ə	LATIN SMALL LETTER REVERSE ITALIC SCHWA
1E004	i	LATIN SMALL LETTER REVERSE ITALIC I
1E005	l	LATIN SMALL LETTER REVERSE ITALIC L
1E006	m	LATIN SMALL LETTER REVERSE ITALIC M
1E007	n	LATIN SMALL LETTER REVERSE ITALIC N
1E008	o	LATIN SMALL LETTER REVERSE ITALIC O
1E009	ø	LATIN SMALL LETTER REVERSE ITALIC O WITH STROKE
1E00A	œ	LATIN SMALL LETTER REVERSE ITALIC OE
1E00B	r	LATIN SMALL LETTER REVERSE ITALIC R
1E00C	s	LATIN SMALL LETTER REVERSE ITALIC S
1E00D	t	LATIN SMALL LETTER REVERSE ITALIC T
1E00E	u	LATIN SMALL LETTER REVERSE ITALIC U
1E00F	y	LATIN SMALL LETTER REVERSE ITALIC Y

Letters for Uralicist dialectology

1E010	b̥	LATIN SMALL LETTER B WITH REVERSED STROKE
1E011	ḋ	LATIN SMALL LETTER D WITH REVERSED STROKE
1E012	ḏ	LATIN SMALL LETTER D WITH HIGH CIRCUMFLEX
1E013	ḡ	LATIN SMALL LETTER G WITH REVERSED STROKE
1E014	ḭ	LATIN SMALL LETTER DOTLESS I WITH STROKE
1E015	ḥ	LATIN SMALL LETTER L WITH HIGH CIRCUMFLEX
1E016	ḷ	LATIN SMALL LETTER L WITH LEFT CURLICUE
1E017	Ḟ	LATIN SMALL LETTER CLOSED OPEN E WITH HORIZONTAL STROKE
1E018	ṭ	LATIN SMALL LETTER T WITH REVERSED CURL
1E019	Ṣ	LATIN SMALL LETTER EZH WITH REVERSED STROKE

Letters for German dialectology

1E020	Ḃ	LATIN SMALL LETTER LIGHTLY CENTRALIZED A
1E021	Ḅ	LATIN SMALL LETTER LIGHTLY CENTRALIZED E
1E022	Ḇ	LATIN SMALL LETTER LIGHTLY CENTRALIZED I
1E023	Ḉ	LATIN SMALL LETTER LIGHTLY CENTRALIZED O
1E024	Ḋ	LATIN SMALL LETTER LIGHTLY CENTRALIZED U
1E025	Ḧ	LATIN SMALL LETTER STRONGLY CENTRALIZED A
1E026	Ḩ	LATIN SMALL LETTER STRONGLY CENTRALIZED E
1E027	Ḫ	LATIN SMALL LETTER STRONGLY CENTRALIZED I
1E028	Ḭ	LATIN SMALL LETTER STRONGLY CENTRALIZED O

1E029	Ṁ	LATIN SMALL LETTER STRONGLY CENTRALIZED U
1E02A	Ṃ	LATIN SMALL LETTER LESS LIGHTLY CENTRALIZED A
1E02B	Ṅ	LATIN SMALL LETTER LESS LIGHTLY CENTRALIZED E
1E02C	Ṇ	LATIN SMALL LETTER LESS LIGHTLY CENTRALIZED I
1E02D	Ṉ	LATIN SMALL LETTER LESS LIGHTLY CENTRALIZED O
1E02E	Ṱ	LATIN SMALL LETTER LESS LIGHTLY CENTRALIZED U
1E02F	Ṛ	LATIN SMALL LETTER LESS STRONGLY CENTRALIZED A
1E030	Ṕ	LATIN SMALL LETTER LESS STRONGLY CENTRALIZED E
1E031	Ṗ	LATIN SMALL LETTER LESS STRONGLY CENTRALIZED I
1E032	Ṙ	LATIN SMALL LETTER LESS STRONGLY CENTRALIZED O
1E033	Ṛ	LATIN SMALL LETTER LESS STRONGLY CENTRALIZED U
1E034	ḃ	LATIN SMALL LETTER B WITH TILDE → 0180 ḃ latin small letter b with stroke
1E035	ḏ	LATIN SMALL LETTER D WITH TILDE → 0111 ḏ latin small letter d with stroke
1E036	f̥	LATIN SMALL LETTER LENIS F → 0066 f̥ latin small letter f
1E037	ḡ	LATIN SMALL LETTER SCRIPT G WITH TILDE
1E038	ḡ	LATIN SMALL LETTER SCRIPT G WITH LOW STROKE → 01E5 ḡ latin small letter g with stroke
1E039	ḥ̈	LATIN SMALL LETTER L WITH DOUBLE MIDDLE TILDE
1E03A	ḥ̈	LATIN SMALL LETTER L WITH PARENTHESES MIDDLE TILDE
1E03B	ḥ̈	LATIN SMALL LETTER L WITH INVERTED LAZY S
1E03C	ḡ	LATIN SMALL LETTER N WITH STROKE
1E03D	ḡ	LATIN SMALL LETTER R WITH SPUR
1E03E	ḡ	LATIN SMALL LETTER STIRRUP R
1E03F	ḡ	LATIN SMALL LETTER SMALL CAPITAL R WITH LONG RIGHT LEG
1E040	ḡ	LATIN SMALL LETTER BROKEN U
1E041	ḡ	LATIN SMALL LETTER U WITH LEFT RING
1E042	ḡ	LATIN SMALL LETTER CHI → 03C7 ḡ greek small letter chi
1E043	ḡ	LATIN SMALL LETTER HALF CHI
1E044	ḡ	LATIN SMALL LETTER CHI WITH LOW LEFT RING
1E045	ḡ	LATIN SMALL LETTER HALF CHI WITH LOW RIGHT RING
1E046	ḡ	LATIN SMALL LETTER X WITH LOW RIGHT RING
1E047	ḡ	LATIN SMALL LETTER CHI WITH LOW LEFT SERIF
1E048	ḡ	LATIN SMALL LETTER HALF CHI WITH LOW LEFT SERIF
1E049	ḡ	LATIN SMALL LETTER STRETCHED O

Letters for Swedish dialectology

1E050	Ḃ	LATIN SMALL LETTER SCRIPT A WITH LOW RING INSIDE
-------	----------	--





1E051	ɑ	LATIN SMALL LETTER SCRIPT A WITH HIGH RING INSIDE	1E06E	ɡ	LATIN SMALL LETTER G WITH DIAGONAL STROKE
1E052	ⱱ	LATIN SMALL LETTER SCRIPT A WITH BACK NOTCH	1E06F	Ɠ	LATIN SMALL LETTER SCRIPT G WITH STROKE
1E053	Ɀ	LATIN SMALL LETTER TURNED SCRIPT A WITH BACK NOTCH	1E070	ɪ	LATIN SMALL LETTER DOTLESS I WITH RIGHT RING
1E054	Ɱ	LATIN SMALL LETTER SCRIPT A WITH NOTCH	1E071	ɿ	LATIN SMALL LETTER DOTLESS I WITH LEFT RING
1E055	Ɱ	LATIN SMALL LETTER TURNED SCRIPT A WITH NOTCH	1E072	ɿ	LATIN SMALL LETTER STRETCHED DOTLESS I → 0237 j latin small letter dotless j
1E056	Ɱ	LATIN SMALL LETTER SCRIPT A TURNED SCRIPT A = coffee bean	1E073	ı	LATIN SMALL LETTER IOTA WITH DOT → 0269 ı latin small letter iota
1E057	Ɱ	LATIN SMALL LETTER SCRIPT A E	1E074	ı	LATIN SMALL LETTER STRETCHED IOTA WITH DOT
1E058	Ɱ	LATIN SMALL LETTER TURNED SCRIPT A E	1E075	к	LATIN SMALL LETTER K WITH CYRILLIC HOOK
1E059	Ɱ	LATIN SMALL LETTER SCRIPT A E WITH NOTCH	1E076	к	LATIN SMALL LETTER K WITH LOW REVERSED CURL palatal k
1E05A	Ɱ	LATIN SMALL LETTER A WITH BACK NOTCH E	1E077	к	LATIN SMALL LETTER K WITH LOW REVERSED CURL AND STROKE
1E05B	Ɱ	LATIN SMALL LETTER TURNED A WITH BACK NOTCH E	1E078	ł	LATIN SMALL LETTER L WITH BARB thick l
1E05C	Ɱ	LATIN SMALL LETTER A WITH BACK NOTCH E WITH NOTCH	1E079	ł	LATIN SMALL LETTER L WITH BARB AND HOOK
1E05D	ɓ	LATIN SMALL LETTER B WITH HIGH CIRCUMFLEX voiceless b	1E07A	ł	LATIN SMALL LETTER L WITH HOOKED BARB voiceless thick l
1E05E	ɓ	LATIN SMALL LETTER FRICATIVE B	1E07B	ł	LATIN SMALL LETTER L WITH HOOKED BARB AND HOOK
1E05F	ɕ	LATIN SMALL LETTER C WITH REVERSED CURL	1E07C	ł	LATIN SMALL LETTER L WITH REVERSED LOW CURL palatal l
1E060	ɕ	LATIN SMALL LETTER D WITH TAIL AND HIGH CIRCUMFLEX voiceless supradental d → 0256 ɕ latin small letter d with tail	1E07D	ɱ	LATIN SMALL LETTER M WITH TILDE SERIF voiceless m
1E061	ɕ	LATIN SMALL LETTER D WITH BARB cacuminal d	1E07E	ɱ	LATIN SMALL LETTER N WITH TILDE SERIF voiceless m
1E062	ɕ	LATIN SMALL LETTER D WITH BARB AND HIGH CIRCUMFLEX voiceless cacuminal d	1E07F	ɱ	LATIN SMALL LETTER N WITH RETROFLEX HOOK AND TILDE SERIF voiceless supradental n → 0273 ɱ latin small letter n with retroflex hook
1E063	ɕ	LATIN SMALL LETTER D WITH REVERSED CURL palatal d	1E080	ɱ	LATIN SMALL LETTER N WITH BARB cacuminal n
1E064	ɕ	LATIN SMALL LETTER D WITH REVERSED CURL AND HIGH CIRCUMFLEX voiceless palatal d	1E081	ɱ	LATIN SMALL LETTER N WITH BARB AND TILDE SERIF voiceless cacuminal n
1E065	ɕ	LATIN SMALL LETTER FRICATIVE D	1E082	ɱ	LATIN SMALL LETTER N WITH LOW REVERSED CURL palatal n
1E066	ɕ	LATIN SMALL LETTER D WITH CYRILLIC HOOK	1E083	ɱ	LATIN SMALL LETTER N WITH LOW REVERSED CURL AND TILDE SERIF voiceless palatal n
1E067	ɕ	LATIN SMALL LETTER ETH WITH STROKE	1E084	ɱ	LATIN SMALL LETTER NWAIR labiodental n
1E068	ɕ	LATIN SMALL LETTER TURNED E WITH NOTCH → 2C78 ɕ latin small letter e with notch	1E085	ɱ	LATIN SMALL LETTER NWAIR WITH TILDE SERIF voiceless labiodental n
1E069	ɕ	LATIN SMALL LETTER STRETCHED E			
1E06A	ɕ	LATIN SMALL LETTER PHI WITH TAIL → 2C77 ɕ latin small letter tailless phi			
1E06B	ɕ	LATIN SMALL LETTER GAMMA WITH RING → 0263 ɕ latin small letter gamma → 03B3 ɕ greek small letter gamma			
1E06C	ɕ	LATIN SMALL LETTER GAMMA WITH VERTICAL TAIL			
1E06D	ɕ	LATIN SMALL LETTER CAROLINGIAN G WITH PALATAL HOOK → 1D83 ɕ latin small letter g with palatal hook			

1E086	ŋ	LATIN SMALL LETTER ENG WITH TILDE SERIF voiceless eng → 014B ŋ latin small letter eng
1E087	ɢ	LATIN SMALL LETTER ENG WITH G BOWL uvular eng
1E088	ɣ	LATIN SMALL LETTER ENG WITH G BOWL AND TILDE SERIF voiceless uvular eng
1E089	ɶ	LATIN SMALL LETTER ENG WITH BACK NOTCH
1E08A	ɷ	LATIN SMALL LETTER ENG WITH WAVY RIGHT LEG
1E08B	ø	LATIN SMALL LETTER REVERSED O WITH STROKE → 01FF ø latin small letter o with stroke
1E08C	o	LATIN SMALL LETTER O WITH LOW CURVED STROKE
1E08D	o	LATIN SMALL LETTER O WITH HIGH CURVED STROKE
1E08E	o	LATIN SMALL LETTER O WITH HIGH OPEN RING egg
1E08F	ø	LATIN SMALL LETTER CLOSED OPEN E WITH STROKE
1E090	ɸ	LATIN SMALL LETTER OU WITH LACHRYMAL TERMINALS
1E091	ɸ	LATIN SMALL LETTER OU WITH LACHRYMAL TERMINALS AND LEFT DIAGONAL STROKE
1E092	ɸ	LATIN SMALL LETTER OU WITH LACHRYMAL TERMINALS AND HIGH OPEN RING
1E093	ɾ	LATIN SMALL LETTER R WITH TILDE SERIF voiceless r
1E094	ɿ	LATIN SMALL LETTER R WITH STROKE
1E095	ɿ	LATIN SMALL LETTER R WITH STROKE AND TILDE SERIF
1E096	ɿ	LATIN SMALL LETTER R WITH DIAGONAL STROKE
1E097	ɿ	LATIN SMALL LETTER SMALL CAPITAL R WITH TILDE SERIF
1E098	ɿ	LATIN SMALL LETTER HALF SCRIPT R
1E099	ɿ	LATIN SMALL LETTER SCRIPT R
1E09A	ɿ	LATIN SMALL LETTER SCRIPT R WITH TILDE SERIF
1E09B	ɿ	LATIN SMALL LETTER TURNED SCRIPT R
1E09C	ɿ	LATIN SMALL LETTER TURNED SCRIPT R WITH TILDE SERIF
1E09D	ɿ	LATIN SMALL LETTER TURNED SCRIPT R WITH STROKE
1E09E	ɿ	LATIN SMALL LETTER TURNED SCRIPT R WITH STROKE AND TILDE SERIF
1E09F	ɿ	LATIN SMALL LETTER S WITH BARB cacuminal s
1E0A0	ɿ	LATIN SMALL LETTER S WITH REVERSED CURL
1E0A1	ɿ	LATIN SMALL LETTER S WITH CYRILLIC HOOK
1E0A2	ɿ	LATIN SMALL LETTER ESJ
1E0A3	ɿ	LATIN SMALL LETTER ESH WITH RING
1E0A4	ɿ	LATIN SMALL LETTER BACK ESH
1E0A5	ɿ	LATIN SMALL LETTER ESH WITH LEFT DIAGONAL STROKE
1E0A6	ɿ	LATIN SMALL LETTER T WITH BARB cacuminal t
1E0A7	ɿ	LATIN SMALL LETTER T WITH LOW REVERSED CURL
1E0A8	ɿ	LATIN SMALL LETTER T WITH CYRILLIC HOOK
1E0A9	u	LATIN SMALL LETTER CLOSED U
1E0AA	u	LATIN SMALL LETTER U WITH NOTCH
1E0AB	u	LATIN SMALL LETTER CLOSED U WITH NOTCH
1E0AC	u	LATIN SMALL LETTER CLOSED TURNED M → 026F u latin small letter turned m
1E0AD	x	LATIN SMALL LETTER X WITH CYRILLIC HOOK → 04FC X cyrillic small letter ha with hook
1E0AE	y	LATIN SMALL LETTER NASAL Y
1E0AF	y	LATIN SMALL LETTER SCRIPT Y
1E0B0	y	LATIN SMALL LETTER SCRIPT Y WITH LEFT RING
1E0B1	y	LATIN SMALL LETTER BROKEN Y delabialized y
1E0B2	y	LATIN SMALL LETTER NASAL SCRIPT Y
1E0B3	y	LATIN SMALL LETTER OPEN Y
1E0B4	y	LATIN SMALL LETTER NASAL OPEN Y
1E0B5	y	LATIN SMALL LETTER SCRIPT Y WITH NOTCH
1E0B6	y	LATIN SMALL LETTER OPEN SCRIPT Y WITH NOTCH
1E0B7	u	LATIN SMALL LETTER U WITH HOOK TAIL → 0265 u latin small letter turned h
1E0B8	z	LATIN SMALL LETTER Z WITH HOOKED BASE
1E0B9	z	LATIN SMALL LETTER Z WITH LOW REVERSED CURL → 0291 z latin small letter z with curl










Letters for the Initial Teaching Alphabet

1E0C0	d	LATIN SMALL LETTER DID
1E0C1	ee	LATIN SMALL LETTER EEN
1E0C2	ie	LATIN SMALL LETTER IDE
1E0C3	ue	LATIN SMALL LETTER UNE
1E0C4	z	LATIN SMALL LETTER ZESS
1E0C5	wh	LATIN SMALL LETTER WHEE
1E0C6	ch	LATIN SMALL LETTER CHAY
1E0C7	th	LATIN SMALL LETTER ITH → 1D7A th latin small letter th with strikethrough
1E0C8	th	LATIN SMALL LETTER THEE
1E0C9	fh	LATIN SMALL LETTER ISH
1E0CA	g	LATIN SMALL LETTER ING → 014B ŋ latin small letter eng
1E0CB	r	LATIN SMALL LETTER ER
1E0CC	a	LATIN SMALL LETTER ASK
1E0CD	au	LATIN SMALL LETTER AUD
1E0CE	oo	LATIN SMALL LETTER OOT
1E0CF	oo	LATIN SMALL LETTER OOD
1E0D0	ou	LATIN SMALL LETTER OUN
1E0D1	oi	LATIN SMALL LETTER OIL

Phonetic letters devised by Benjamin Franklin

- 1E0E0  LATIN SMALL LETTER SCRIPT A WITH NOTCH AND HIGH NOTCH
- 1E0E1  LATIN SMALL LETTER H WITH RETROFLEX HOOK AND ANGLED ASCENDER
- 1E0E2  LATIN SMALL LETTER H WITH LONG LEFT LEG AND CROSSED ASCENDER
- 1E0E3  LATIN SMALL LETTER H WITH LONG RIGHT LEG

Modifier letters

- 1E0F0  MODIFIER LETTER SMALL AE
- 1E0F1  MODIFIER LETTER SMALL SCRIPT A E
- 1E0F2  MODIFIER LETTER SMALL SCHWA WITH HOOK
- 1E0F3  MODIFIER LETTER SMALL I
- 1E0F4  MODIFIER LETTER SMALL DOTLESS I
- 1E0F5  MODIFIER LETTER SMALL CAPITAL R
- 1E0F6  MODIFIER LETTER SMALL SCRIPT R
- 1E0F7  MODIFIER LETTER SMALL REVERSE ITALIC S
- 1E0F8  MODIFIER LETTER REVERSED PRIME

A. Administrative

1. Title

Exploratory proposal to encode Germanicist, Nordicist, and other phonetic characters in the UCS

2. Requester's name

Michael Everson

3. Requester type (Member body/Liaison/Individual contribution)

Individual contribution.

4. Submission date

2008-11-27

5. Requester's reference (if applicable)

6. Choose one of the following:

6a. This is a complete proposal

No.

6b. More information will be provided later

Yes.

B. Technical – General

1. Choose one of the following:

1a. This proposal is for a new script (set of characters)

No.

1b. Proposed name of script

1c. The proposal is for addition of character(s) to an existing block

Yes.

1d. Name of the existing block

Combining Diacritical Marks Supplement, Combining Diacritical Marks Extended, Phonetic Extensions-A

2. Number of characters in proposal

241.

3. Proposed category (A-Contemporary; B.1-Specialized (small collection); B.2-Specialized (large collection); C-Major extinct; D-Attested extinct; E-Minor extinct; F-Archaic Hieroglyphic or Ideographic; G-Obscure or questionable usage symbols)

Category B.2.

4a. Is a repertoire including character names provided?

Yes.

4b. If YES, are the names in accordance with the “character naming guidelines” in Annex L of P&P document?

Yes.

4c. Are the character shapes attached in a legible form suitable for review?

Yes.

5a. Who will provide the appropriate computerized font (ordered preference: True Type, or PostScript format) for publishing the standard?

Michael Everson.

5b. If available now, identify source(s) for the font (include address, e-mail, ftp-site, etc.) and indicate the tools used:

Michael Everson, Fontographer.

6a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided?

Yes.

6b. Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed characters attached?

No.

7. Does the proposal address other aspects of character data processing (if applicable) such as input, presentation, sorting, searching, indexing, transliteration etc. (if yes please enclose information)?

No.

8. Submitters are invited to provide any additional information about Properties of the proposed Character(s) or Script that will assist in correct understanding of and correct linguistic processing of the proposed character(s) or script. Examples of such properties are: Casing information, Numeric information, Currency information, Display behaviour information such as line breaks, widths etc., Combining behaviour, Spacing behaviour, Directional behaviour, Default Collation behaviour, relevance in Mark Up contexts, Compatibility equivalence and other Unicode normalization related information. See the Unicode standard at <http://www.unicode.org> for such information on other scripts. Also see Unicode Character Database <http://www.unicode.org/Public/UNIDATA/UnicodeCharacterDatabase.html> and associated Unicode Technical Reports for information needed for consideration by the Unicode Technical Committee for inclusion in the Unicode Standard.

See above.

C. Technical – Justification

1. Has this proposal for addition of character(s) been submitted before? If YES, explain.

No.

2a. Has contact been made to members of the user community (for example: National Body, user groups of the script or characters, other experts, etc.)?

Yes.

2b. If YES, with whom?

Odd Einar Haugen, Janne Johannessen, Ulrike Kramer, Marc Wilhelm Küster, Klaas Ruppel, Trond Trosterud, Karl Pentzlin, Sven Peters, Henrik Rosenkvist, Øystein Vangsnes, Eveline Wandl-Vogt, Peter Weiss.

2c. If YES, available relevant documents

3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included?

Germanicists, Nordicists, Anglicists, Uralicists.

4a. The context of use for the proposed characters (type of use; common or rare)

Used historically and in modern editions.

4b. Reference

5a. Are the proposed characters in current use by the user community?

Yes.

5b. If YES, where?

Scholarly publications.

6a. After giving due considerations to the principles in the P&P document must the proposed characters be entirely in the BMP?

No.

6b. If YES, is a rationale provided?

6c. If YES, reference

7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)?

No.

8a. Can any of the proposed characters be considered a presentation form of an existing character or character sequence?

No.

8b. If YES, is a rationale for its inclusion provided?

8c. If YES, reference

9a. Can any of the proposed characters be encoded using a composed character sequence of either existing characters or other proposed characters?

No.

9b. If YES, is a rationale for its inclusion provided?

9c. If YES, reference

10a. Can any of the proposed character(s) be considered to be similar (in appearance or function) to an existing character?

No, surely not!

10b. If YES, is a rationale for its inclusion provided?

10c. If YES, reference

11a. Does the proposal include use of combining characters and/or use of composite sequences (see clauses 4.12 and 4.14 in ISO/IEC 10646-1: 2000)?

Yes.

11b. If YES, is a rationale for such use provided?

No.

11c. If YES, reference

11d. Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided?

No.

11e. If YES, reference

12a. Does the proposal contain characters with any special properties such as control function or similar semantics?

No.

12b. If YES, describe in detail (include attachment if necessary)

13a. Does the proposal contain any Ideographic compatibility character(s)?

No.

13b. If YES, is the equivalent corresponding unified ideographic character(s) identified?